



## Clitiques et constructions topicalisées dans une grammaire GPSG du français

Gabriel G. Bès

### ► To cite this version:

Gabriel G. Bès. Clitiques et constructions topicalisées dans une grammaire GPSG du français. *Lexique*, 1988, *Lexique et paraphrase*, 6, pp.55-81. hal-01096803

**HAL Id: hal-01096803**

**<https://hal.science/hal-01096803>**

Submitted on 18 Dec 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Clitiques et constructions topicalisées dans une grammaire *GPSG* du français\*

Gabriel G. Bès

Il n'est pas improbable que 1982 soit un jour considérée comme une année charnière dans l'histoire de la linguistique : *Lectures on government and binding* de Noam Chomsky vient de paraître et, grâce à beaucoup de facteurs, parmi lesquels le GLOW, *GB* occupe le devant de la scène dans de vastes secteurs du monde académique, et ceci particulièrement en Europe. Mais 1982 voit aussi la publication des travaux de Joan Bresnan et de Gerald Gazdar qui, eux, notamment en France, passent beaucoup plus inaperçus. Est-ce que 1982, un quart de siècle plus tard, ne reproduit pas, mais en des termes inversés, la situation notée par Bach (1965) pour 1957, année qui avait vu la publication de l'anthologie de Martin Joos *Readings in Structural Linguistics*, expression de la pensée dominante de l'académie, alors que dans *Syntactic Structures*, la même année, Chomsky en remettait en cause les fondements ?

Un quart de siècle dominé par la pensée chomskyenne. Le bilan — à faire — n'est pas simple. Vraisemblablement, on ne dira pas assez tout ce que la linguistique doit à Chomsky. Mais ses derniers modèles, mal explicités, surabondants en structures de types différents, avec un alignement de boîtes noires construites en fonction de contraintes d'entrée et sortie sans que les opérations internes soient dûment caractérisées et avec des représentations sémantiques aux capacités formelles douteuses, conduisent à la perplexité. Et si le souci est de mettre en rapport grammaires linguistiques et analyseurs automatiques, la méfiance à l'égard des modèles chomskyens ne peut qu'augmenter.

Un quart de siècle qui a vu s'installer une idée dominante : les règles de réécriture indépendantes du contexte ne sont pas un outil suffisamment puissant pour rendre compte des langues naturelles. D'où la nécessité de plus d'une structure pour représenter un énoncé ; d'où la nécessité d'opérations permettant d'associer ces structures et, en particulier, les opérations de transformation.

Avec des modalités différentes, Bresnan et Gazdar, dans leurs publications de 1982 — travaux qui vont très au-delà d'une première esquisse

\* Ce travail a été partiellement accompli dans le cadre du Projet ESPRIT 393 ACORD. Le paragraphe 2 reprend, en la modifiant et l'élargissant, une analyse préalablement présentée dans Bès (1986).

d'une idée féconde — proposent des modèles sans transformations. Gazdar explore jusqu'au bout la simplification du modèle ; des règles syntagmatiques indépendantes du contexte et un seul niveau de représentation suffisent pour décrire et expliquer les langues naturelles<sup>1</sup>.

Trois années plus tard est publié *Generalized Phrase Structure Grammar* (désormais *GPSG*), par G. Gazdar, E. Klein, G. Pullum et I. Sag. Le modèle est présenté de manière rigoureuse, claire, systématique, avec sa syntaxe et sa sémantique. Il y est appliqué à l'anglais, où une grande variété de constructions sont couvertes (interrogatives, passives, relatives, enchâssées (infinitives et finies), clivées, dépendances arbitrairement distantes, coordination, ...). On y voit leur syntaxe et leur sémantique se construire de manière limpide, sans l'introduction miraculeuse, pour faire tourner le modèle, de principes *ad hoc* au fond des notes, sans le refrain « les détails mis à part ». Passer du cauchemar de la rhétorique chomskyenne au texte de *GPSG* est un vrai plaisir intellectuel.

Il est certes vrai que tout n'y fonctionne pas correctement et on peut même s'étonner des négligences sur des points importants<sup>2</sup>. Mais l'immense avantage des écrits linguistiques orientés par des personnalités telles que Bresnan, Gazdar, Klein, Bach, Dowty, Sag et d'autres, c'est justement qu'ils sont faits plus pour montrer et convaincre que pour cacher et s'imposer. Le formalisme de *GPSG* peut donc être manipulé sans crainte excessive et confronté à des problèmes descriptifs spécifiques.

Voici donc l'objectif de ce travail. Il a semblé intéressant de présenter *GPSG*, un des modèles postchomskyens qui bénéficient d'une manière de raisonner chomskyenne (description et explication de la compétence, recherche de généralisations), mais qui ne fait aucune concession au brouillard rhétorique. Avec seulement des règles syntagmatiques, le modèle peut traiter les dépendances arbitrairement distantes et une variété de phénomènes (permutations, extractions, accord, etc.) qui, jusqu'en 1982, avaient été considérés comme des raisons incontestées pour exiger des transformations. Un des prix à payer est relativement lourd : les entrées se multiplient dans le lexique et, ce qui paraît plus important, l'information qu'on doit y associer est très importante. Dans ce qui suit, le paragraphe 1 est consacré à une présentation rapide du modèle. Celui-ci, dans le paragraphe 2, est appliqué à une gamme (non exhaustive) d'observations concernant les clitiques et les constructions topicalisées en français.

*GPSG* bénéficie aussi des apports de la grammaire de Montague. Les arbres syntaxiques spécifiés par une grammaire *GPSG* sont associés à une traduction sémantique en logique intensionnelle, elle-même interprétable dans un univers donné. A partir de la solution syntaxique proposée dans le paragraphe 2, on montre, dans le paragraphe 3, comment *GPSG* peut traiter certaines paraphrases : il va associer des traductions sémantiques formellement équivalentes à des arbres syntaxiques différents.

Le paragraphe 4 enfin est consacré à tenter une évaluation des possibilités du modèle à partir des problèmes traités.

## 1. GPSG

Les grammaires particulières spécifiées par GPSG constituent un sous-ensemble des grammaires indépendantes du contexte ; elles n'incorporent pas de transformations et les représentations correspondantes possèdent un seul niveau syntaxique. Les règles opèrent sur des catégories, qui sont des ensembles de traits.

Une grammaire particulière doit se spécifier par :

- (1) a Enumération de traits avec leurs valeurs possibles
- b Restrictions sur la cooccurrence des traits (= FCR, *feature co-occurrence restrictions*)
- c Spécification des traits par défaut (= FSD, *feature specification defaults*)
- d Règles de dominance immédiate (= règles ID, *immediate dominance rules*)
- e Métarègles (= M, *metarules*)
- f Stipulations de précedence linéaire (= stipulations LP, *linear precedence statements*)
- g Entrées lexicales (= EL)

Des principes généraux — dits principes universaux d'instanciation — vont régler les rapports entre les arbres — représentations associées aux énoncés — et les différents éléments de la grammaire. On présentera d'abord ceux-ci pour revenir par la suite à ceux-là<sup>3</sup>.

Un trait est constitué d'une étiquette associée à des valeurs : il peut être noté <ETIQUETTE ; VALEURS>. Une catégorie est un ensemble de traits ou, plus précisément, une fonction partielle qui va associer l'ensemble des étiquettes à l'ensemble des valeurs. Dans une catégorie chaque étiquette ne sera associée qu'à une seule valeur. Parmi les traits, il y en a à valeurs atomiques (par exemple <PLURIEL ; ± >, <PERSONNE ; 1, 2, 3 > et d'autres à valeur catégorielle : ceux-ci sont des traits qui ont des catégories comme valeurs. Les traits sont catégorisés en traits FOOT et traits HEAD.

Les FCR et les FSD sont, formellement, des implications ; ils vont exprimer le fait que la présence d'un trait dans une catégorie entraîne nécessairement la présence ou l'absence d'autres traits dans la même catégorie. Un FCR est une exigence qui porte sur les cas explicitement couverts dans la formulation ; par exemple, le FCR [A +] ⊃ [B -] dit que toute catégorie constituée par le trait A avec valeur + comporte aussi le trait B avec valeur -. Les FSD sont aussi des exigences sur la cooccurrence des traits dans les catégories, mais elles sont valables pour tous les cas où la grammaire n'exprime pas explicitement le contraire. Un FSD [A +] s'interprète donc : dans tous les cas où la grammaire n'assigne pas explicitement une valeur différente au trait A, celui-ci sera +.

Les règles ID opèrent sur des catégories. Elles n'expriment que des relations de dominance entre la catégorie mère (notée désormais Co) et la ou les catégories filles (Ci... Cn), dont l'une, notée H (= tête), aura des propriétés particulières. Les filles ne sont pas ordonnées entre elles, de telle manière que la règle (2) spécifie les arbres locaux de (3)

ou, en notation équivalente, de (4) (sur la notation  $X'$ , cf. ci-dessous).

$$(2) \quad Co \rightarrow H, Ci, Cj$$

$$(3) \quad (a) \begin{array}{c} Co' \\ / \quad | \quad \backslash \\ H' \quad Ci' \quad Cj' \end{array}$$

$$(b) \begin{array}{c} Co' \\ / \quad | \quad \backslash \\ Ci' \quad H' \quad Cj' \end{array}$$

$$(c) \begin{array}{c} Co' \\ / \quad | \quad \backslash \\ Cj' \quad Ci' \quad H' \end{array}$$

$$(4) \quad (a) \begin{array}{c} Co' \\ \quad H' \\ \quad \quad Ci' \\ \quad \quad \quad Cj' \end{array}$$

$$(b) \begin{array}{c} Co' \\ \quad \quad Ci' \\ \quad \quad \quad H' \\ \quad \quad \quad \quad Cj' \end{array}$$

$$(c) \begin{array}{c} Co' \\ \quad \quad \quad Cj' \\ \quad \quad \quad \quad Ci' \\ \quad \quad \quad \quad \quad H' \end{array}$$

Une règle ID lexicale est une règle ID dont la tête va dominer un symbole terminal.

Les arbres complexes sont le résultat de l'imbrication d'arbres locaux simples ; chacun de ceux-ci, constitué d'une mère et d'une ou plusieurs filles, étant associé à une règle ID différente. Par exemple, (5) se décompose dans les arbres locaux (6) et (7), spécifiés respectivement par les règles (8) et (9). On peut observer que la catégorie VP est fille dans (6) mais mère dans (7).

$$(5) \begin{array}{c} S \\ / \quad \backslash \\ NP \quad VP \\ \quad \quad / \quad \backslash \\ \quad \quad H \quad NP \end{array}$$

$$(6) \begin{array}{c} S \\ / \quad \backslash \\ NP \quad VP \end{array}$$

$$(7) \begin{array}{c} VP \\ / \quad \backslash \\ H \quad NP \end{array}$$

$$(8) \quad S \rightarrow NP, VP$$

$$(9) \quad VP \rightarrow H, NP$$

Les métarègles opèrent sur les règles ID lexicales et seulement sur celles-ci. Elles permettent d'obtenir des règles ID lexicales à partir des règles ID lexicales.

Les stipulations de précedence linéaire vont caractériser l'ordre des catégories filles dans les arbres. Avec la stipulation LP de (10), par exemple, l'arbre (3c) (ou (4c)) ne serait pas admis comme bien formé (le symbole « < » doit se lire « précède »):

$$(10) \quad Ci < Cj$$

La relation entre arbres locaux et règles, que celles-ci soient listées ou introduites par métarègle, est conditionnée par les principes universaux d'instanciation. La fonction de projection associe les catégories dans les arbres locaux (ou catégories instanciées) aux catégories dans les règles (ou catégories licenciatrices). On aura ainsi, en général,  $\emptyset(C) = C'$ , où  $C$  est une catégorie dans une règle et  $C'$  une catégorie

dans un arbre local. La fonction  $\emptyset$  associe ainsi chaque catégorie dans une règle à chaque catégorie dans un arbre (cf. ci-dessus la relation de (2) avec (3) ou (4)). D'un trait présent dans C et dans C', on dira qu'il est *licencié*; d'un trait attesté dans C' et absent de C, on dira qu'il est *instancié dans C'*.

L'extension et l'unification sont les deux rapports qui sont à la base des relations entre catégories. Ils peuvent conceptuellement s'exprimer comme dans (11) et (12):

- (11) C'est une extension de C si toutes les valeurs de C sont aussi dans C'.
- (12) Co' présente une unification par rapport à ses catégories filles si Co' est une extension de chacune de ses catégories filles.

La stratégie générale du modèle consiste à définir des exigences qui vont porter sur les catégories associées par la fonction de projection ou sur les catégories dans les arbres locaux. Ces exigences vont demander la présence obligatoire des traits dans les deux termes associés par une projection ou dans les catégories d'un arbre local qui sont dans un rapport déterminé. Elles seront modulées selon:

- (a) le type de trait (d'où l'importance de la catégorisation des traits en FOOT, HEAD, atomiques, etc.);
- (b) le type de relation (extension ou unification, cf. ci-dessus) entre catégories;
- (c) l'élément de la grammaire qui doit être satisfait (FCR, FSD, etc.);
- (d) des principes d'instanciation spécifiques (pour l'accord, les traits FOOT, etc.).

Ces exigences s'additionnent: lorsque toutes sont satisfaites, il est possible de définir la notion d'*arbre admissible par rapport à l'ensemble des règles ID d'une grammaire*. Elles sont (entre parenthèses, l'élément de la grammaire ou le problème linguistique auquel chaque exigence se rattache):

- (a) projection légale (FCR);
- (b) principe FFP (*Foot feature principle*, traits FOOT);
- (c) principe CAP (*Control agreement principle*, accord et contrôle des verbes à l'infinitif);
- (d) principe HFC (*Head feature convention*, transmission des traits par les catégories têtes);
- (e) acceptabilité LP (stipulations LP);
- (f) admissibilité (FSD). On examinera par la suite individuellement chacune de ces exigences.

*Projection légale.* On dira d'un arbre t qu'il est une projection légale d'une règle r ssi:

- (a) dans toute projection  $\emptyset(C) = C'$ , C' est une extension de C;
- (b) les FCR sont satisfaits dans C'.

Par exemple, (15) est une projection légale de (13), car chacune des catégories de (15) est une extension des catégories licenciatrices correspondantes de (13), et (14) y est respecté. En revanche, bien que les catégories de (16) soient des extensions des catégories respectives de (13),

(16) n'est pas une projection légale de (13), car le FCR (14) n'y est pas respecté (dans NP[FIN]) :

(13) VP → H[3], NP

(14) [PFORM] ⊃ [+V, -N]

(15) VP[FIN]  
       H[3, FIN]  
       NP[PLU+]

(16) VP  
       H[3]  
       NP[FIN]

*Principe FFP.* Les traits FOOT instanciés dans la mère doivent présenter une unification par rapport aux traits FOOT instanciés dans les filles.

Par exemple, en relation à (17), l'arbre local (18) respecte le principe FFP, alors que (19) ne le fait pas (le trait SLASH, noté “/”, qui est à valeur catégorielle, est un trait FOOT) :

(17) VP → V[n], NP, PP

(18) VP/NP  
       NP/NP  
       PP

(19) VP  
       NP/NP  
       PP

Le principe FFP et les traits FOOT, en particulier SLASH, sont cruciaux pour rendre compte des dépendances arbitrairement éloignées : cf. ci-dessous leur utilisation au § 2.3. *in fine*. Intuitivement, le trait /x, dans C/x, introduit l'information : le constituant x manque dans la catégorie C. Le principe FFP permet de transmettre cette information à travers les arbres locaux. Il faut noter que FFP ne porte que sur les traits FOOT instanciés ; le SLASH licencié, en particulier, aura un comportement différent et pour la syntaxe et pour la sémantique.

*Principe CAP.* Le principe CAP règle les questions d'accord morphologique et du contrôle des verbes à l'infinitif. Il opère sur les traits à valeur catégorielle AGR (= agreement) et SLASH, et ceci, ce qui semble très remarquable dans le modèle, en fonction des types sémantiques associés aux constituants dans la sémantique de GPSG (cf. ci-dessous au § 3 les éléments de base), où on distingue entre foncteurs et arguments. Pour les cas d'accord morphologique, CAP exigera que la valeur du trait AGR de l'argument soit reproduite dans le foncteur. Ceci permettra de transmettre les traits morphologiques d'accord dans, par exemple, les arbres locaux de (19) et (20) :

(19) S  
       NP[AGR α]  
       VP[AGR α]

- (20) NP  
       Det[AGR  $\beta$ ]  
       N' [AGR  $\beta$ ]

où VP et Det sont les foncteurs et NP et N' les arguments. Pour satisfaire CAP, le NP de (19) (qui correspond au sujet), doit posséder la même valeur catégorielle pour AGR que le VP. *Mutatis mutandis*, la même exigence vaut pour la tête nominale associée à N' (argument) et le Det respectif (sur l'utilisation de CAP dans le contrôle des infinitives, cf. dans ce même volume l'article de K. Baschung).

*Principe HFC.* Ce principe règle la transmission de l'information entre une catégorie mère et la tête (H), qui est une catégorie distinguée parmi les filles. Il concerne les traits HEAD et, intuitivement, est l'héritier de certains aspects de la convention X barre : les traits HEAD — et, donc,  $\langle V; \pm \rangle$  et  $\langle N; \pm \rangle$  — seront transmis par HFC de mère en tête, ce qui va assurer, parmi d'autres effets, que les catégories S, VP et V soient toutes  $[+V, -N]$ . Le principe HEAD peut conceptuellement s'exprimer par (21) et (22) :

- (21) Traits libres : ensemble de traits qui sont susceptibles d'être instanciés après satisfaction des FCR, de FFP et de CAP.  
 (22) Si un trait est HEAD et est libre, et qu'il est instancié dans la mère, il doit être instancié dans la tête et *vice versa*.

Par exemple, si l'on a la règle ID (23) et le FCR (24)

- (23) VP  $\rightarrow$  H[n], NP  
 (24) [BAR 0]  $\equiv$  [N] & [SUBCAT] & [V]

H[n]', dans un arbre local sera [n, BAR 0, +V, -N]. En effet, la satisfaction des FCR, FFP et CAP n'y entraîne pas l'absence de  $[+V, -N]$  ; ceux-ci sont donc libres. Comme ils sont HEAD, (22) exige qu'ils y soient instanciés. En revanche, bien que  $\langle \text{BAR} ; 0, 1, 2 \rangle$  soit un trait HEAD, (24) exige [BAR 0] dans H[n] ; [BAR 2] de VP n'est donc pas libre et HFC ne le transmet pas sur la tête.

*Acceptabilité LP.* Un arbre local, projection de la règle r, dans lequel Ci' précède Cj', est acceptable LP ssi : (a) il n'y a pas de stipulation LP telle que  $K_j < K_i$  ; (b) Ci' et Cj' sont respectivement des extensions de  $K_i$  et de  $K_j$ .

Une conséquence de cette définition est la possibilité, à partir d'une stipulation comme dans (25), d'obtenir l'interdiction de (26), Ca' étant une extension de  $K_a$  :

- (25)  $K_a < K_a$   
 (26) \*Ca', Ca'

*Admissibilité.* Ce principe d'instanciation va régler le rapport des arbres locaux avec les FSD, tout en récapitulant les exigences précédentes. Il peut se caractériser moyennant (27) et (28) :

- (27) Un arbre local qui satisfait FCR, FFP, CAP et HFC, et qui est acceptable LP, est une projection candidate à l'admissibilité d'une règle r.



- (28) Une projection candidate à l'admissibilité de  $r$  est une projection admissible de  $r$  ssi :

Les FSD  $y$  sont satisfaits

ou bien

Il n'existe pas de projection candidate à l'admissibilité de  $r$  où les FSD soient respectés

ou bien

La satisfaction des FSD entraînerait une modification d'une autre catégorie dans la projection de  $r$ .

Parmi les arbres locaux il reste à introduire le cas particulier d'une catégorie mère qui domine une seule fille, celle-ci étant une entrée lexicale (= EL) :

- (29)  $Co$   
entrée lexicale<sub>i</sub>

Une EL possède quatre types d'information :

- (a) une suite phonologique ;
- (b) une catégorie (c'est-à-dire un ensemble de traits, objet du même type que les catégories dans les arbres et les règles) ;
- (c) des indications sur les irrégularités morphologiques ;
- (d) une traduction sémantique.

Les points (a) et (c) ne sont pas traités dans *GPSG* ; sur d) cf. ci-dessous § 3.

On dira de (29) qu'il est un *arbre local terminé*. On exige que  $Co$  y soit une extension de la catégorie de l'entrée lexicale. Un arbre est un *arbre terminé* si ses symboles terminaux sont des EL.

A partir des définitions précédentes, il est possible d'introduire la caractérisation d'*admissibilité d'un arbre par rapport à l'ensemble des règles ID d'une grammaire* :

- (30) Un arbre  $t$  est admissible par un ensemble  $R$  des règles ID ssi : (a) il est terminé ; (b) tout sous-arbre local de  $t$  soit est un arbre local terminé soit est une projection admissible d'une règle  $r$  de  $R$ .

## 2. SYNTAXE

Le paragraphe 2.1. synthétise les observations dont la solution proposée par la suite (paragraphe 2.2.) devrait rendre compte. Les commentaires du paragraphe 2.3. mettent systématiquement en relation celle-ci avec celles-là.

### 2.1. *Le champ observationnel visé*

On n'abordera, parmi les structures interrogatives du français, que l'inversion pronominale du sujet dans les phrases racines :

- (1) Regardes-tu ce livre ?
- (2) Regarde-t-il ce livre ?
- (3) \*Regarde Jacques ce livre ?

Les clitiques peuvent être intuitivement classés comme des PRO-formes. Tous les clitiques ne sont pas des PRO-noms (comme l'est celui de (4)) ; il y a des PRO-adjectifs, des PRO-syntagmes prépositionnels, des PRO-syntagmes verbaux et des PRO-phrases, comme dans respectivement, (5) à (8) :

- (4) Pierre, Marie l'aime
- (5) Intelligente, Marie l'est
- (6) A la poste, Pierre y va souvent
- (7) Apprendre les mathématiques, Pierre le souhaite
- (8) Qu'il est déjà tard, Pierre le sait

L'emplacement des clitiques obéit à des contraintes précises. On abordera ici les clitiques dans les phrases affirmatives ou dans les interrogatives de type (1) à (2). Ils sont situés à gauche d'une forme verbale ; celle-ci, dans le champ observé, sera toujours une forme verbale finie simple. Tout clitique peut se situer seul à la gauche du verbe. Les transitions possibles entre eux peuvent se condenser dans la matrice suivante<sup>4</sup> :

	se(ac)	se(dat)	te(ac)	te(dat)	le	lui	y	en(ac)	en(de)
se(ac)	*	*	*	*	*	*	G	*	G
se(dat)	*	*	*	*	G	*	*	G	G
te(ac)	*	*	*	*	*	*	G	*	G
te(dat)	*	*	*	*	G	*	*	G	G
le	*	*	*	*	*	G	G	*	G
lui	*	*	*	*	*	*	*	G	G
y	*	*	*	*	*	*	*	G	*
en(ac)	*	*	*	*	*	*	*	*	*
en(de)	*	*	*	*	*	*	*	*	*

Dans une transition entre clitiques précédant une forme verbale, le clitique de gauche est représenté, dans (9), en colonne, et le clitique de droite sur la ligne horizontale. (9) exprime ainsi le fait que, par exemple, *le* peut être suivi de *lui*, mais non de *te* :

- (10) Marie le lui donne
- (11) \*Marie le te donne.

Il est ainsi bien établi que l'ordre entre clitiques obéit à des contraintes spécifiques, très différentes de celles qui règlent l'ordre entre constituants lexicaux (comparer (10) et (11) avec (12) et (13)), malgré le fait qu'entre ceux-ci et ceux-là, il y ait des rapports systématiques clairs (cf. les paires de (14)) :

- (12) Marie donne un livre à Pierre
- (13) Marie donne à Pierre un livre
- (14) a Pierre pense à la promenade
- a' Pierre y pense
- b Pierre raffole de cela
- b' Pierre en raffole
- c Pierre donne des fleurs à Marie
- c' Pierre en donne à Marie

La représentation phonique des clitiques diffère de celle des formes non clitiques :

- (15) a Toi, Marie te connaît  
 b Marie pense à toi  
 c Tu aimes Marie  
 d \*Te, Marie tu connaît  
 e \*Marie pense à te.

Immédiatement à gauche de la phrase, dans une position qui sera appelée ici *position topique*, on trouve une variété de constructions (cf. (4) à (8)). Elles se combinent aussi avec les interrogatives (1) et (2) :

- (16) Toi, regardes-tu ce livre ?  
 (17) Pierre, aime-t-il ce livre ?

Dans cette même position on trouve aussi des constituants interrogés :

- (18) Quel livre connais-tu ?

Les rapports entre les constructions dans la position topique et les constituants de la phrase (y compris le sujet), peuvent se résumer dans (19) et (20), respectivement illustrés dans (21) et (22) :

- (19) Une construction non interrogative dans la position topique doit être associée à une PRO-forme ouverte (c'est-à-dire, avec représentation phonique), clitique ou non clitique, dans la phrase.  
 (20) Une construction interrogative dans la position topique ne doit pas être associée à une PRO-forme ouverte.  
 (21) a A Pierre, Marie lui donne un livre  
 a' \*A Pierre, Marie donne un livre à Jacques  
 b Pierre, aime-t-il les tulipes ?  
 b' \*Marie, aime-t-il les tulipes ?  
 (22) a Quel livre aime-t-il ?  
 a' \*Quel livre l'aime-t-il ?

## 2.2. Solution proposée

On présentera par la suite les spécifications d'une grammaire *GPSG* du français qui devrait être susceptible de rendre compte du champ observationnel visé dans le paragraphe précédent.

### 2.2.1. Traits et valeurs

#### Traits HEAD

Traits :	Valeurs :
AGR	CAT
BAR(re)	{ 0, 1, 2 }
CAS	{ NOM(inatif), AC(cusatif), OBL(ique) }
CLBL(cliticisable)	{ +, - }

CLIT(ique)	{ +, - }
DET(erminant)	{ DEF(ini), IND(éfini), INT(errogatif) }
GEN(re)	{ MASC(ulin), FEM(inin) }
HUM(ain)	{ +, - }
INV(ersion)	{ +, - }
N	{ +, - }
PER(sonne)	{ 1, 2, 3 }
PFORM(e)	{ à, de, en, avec }
PROF(orme)	{ PRON(om), PROV(verb.), PROA(djectif), PROS(phrase) }
PLU(riel)	{ +, - }
REF(érentiel)	{ +, - }
REFL(exif)	{ +, - }
SLASH	CAT
SUBCAT(égorisation)	{ 1, ..., n } ∪ { que, ... }
SUJ(et)	{ +, - }
TEM(ps)	{ IMP, PRES, ... }
V	{ +, - }
VFORM(e)	{ FIN, INF... }

*Traits FOOT*

Traits :	Valeurs :
SLASH	CAT
WH	CAT

*Traits ni FOOT ni HEAD*

Traits :	Valeurs :
NUL	{ +, - }
WHMOR	{ Q, ... }
COMP	{ que, ... }

## 2.2.2. FCR

FCR 1	: [PRON, PER 1]	⊃	[+REF, +HUM]
FCR 2	: [PRON, PER 2]	⊃	[+REF, +HUM]
FCR 3	: [NOM] V [OBL]	⊃	[-CLIT]
FCR 4	: [+CLIT]	⊃	[PROF]
FCR 5	: [PROF] & [NP]	⊃	[PRON]
FCR 6	: [PROF] & [VP]	⊃	[+CLIT, PROV]
FCR 7	: [PROF] & [S]	⊃	[+CLIT, PROS]
FCR 8	: [PROF] & [AP]	⊃	[+CLIT, PROA]
FCR 9	: [+CLIT, PRON]	⊃	REF
FCR 10	: [CLBL +, -CLIT]	⊃ ~	[PROF]
FCR 11	: [CLBL -]	⊃	[-CLIT]
FCR 12	: [+REFL]	⊃	[+CLIT, +REF, PRON, +HUM, PER 3]
FCR 13	: [PRON]	⊃	[PER] & [+N, -V]
FCR 14	: [PFORM]	⊃	[NP] V [PP]
FCR 15	: [INT]	⊃	[+Q]
FCR 16	: [+INV]	⊃	[FIN]
FCR 17	: [S/NP[PRON, +REF]]	⊃	[S/NP[PRON, +REF, OBL]]
FCR 18	: [S/SX[+Q]]	⊃	[+INV]
FCR 19	: [VFORM]	⊃	[+V, -N]

FCR 20	:	[BAR 0] $\equiv$ [N] & [SUBCAT] & [V]	
FCR 21	:	[BAR 1] $\setminus$	$\supset \sim$ [SUBCAT]
FCR 22	:	[BAR 2] $-$	$\supset \sim$ [SUBCAT]
FCR 23	:	[+INV, BAR 2]	$\supset$ [+SUJ]
FCR 24	:	[+SUJ]	$\supset$ [+V, -N, BAR 2]
FCR 25	:	[AGR]	$\supset$ [-N, +V]
FCR 26	:	[FIN, AGR NP]	$\supset$ [AGR NP[NOM]]
FCR 27	:	[+NUL]	[SLASH]
FCR 28	:	[+V, -N, BAR 0, $\sim$ PROF]	$\supset$ [-CLIT]

### 2.2.3. FSD

FSD 1	:	[FIN]	
FSD 2	:	$\sim$ [NUL]	
FSD 3	:	$\sim$ [CAS]	
FSD 4	:	[PFORM]	$\supset$ [BAR 0]
FSD 5	:	$\sim$ [REF]	
FSD 6	:	[+N, -V]	$\supset \sim$ [PFORM]
FSD 7	:	$\sim$ [/ [SX, $\sim$ Q]]	
FSD 8	:	[NP[AC V PFORM $\alpha$ ]]	$\supset$ [CLBL+]

### 2.2.4. Règles ID

R 1	:	S $\rightarrow$ NP, H[-SUJ]	
R 2	:	S[COMP $\alpha$ ] $\rightarrow$ {[SUBCAT $\alpha$ ]}, H[COMP NIL, -Q] où $\alpha$ est dans { <i>que</i> ,...}	
R 3	:	S $\rightarrow$ X <sup>2</sup> , H/X <sup>2</sup>	
R 4	:	AP $\rightarrow$ H <sup>1</sup>	
R 5	:	PP $\rightarrow$ P <sup>1</sup>	
R 6	:	NP[DET $\alpha$ ] $\rightarrow$ {[SUBCAT $\alpha$ ]}, H <sup>1</sup> où $\alpha$ est dans {DEF, IND, INT}	
R 7	:	NP[à, -CLIT] $\rightarrow$ P[1], H[ PFORM, OBL] (à)	
R 8	:	NP[de, -CLIT] $\rightarrow$ P[2], H[ PFORM, OBL] (de)	
R 9	:	NP[avec, -CLIT] $\rightarrow$ P[3], H[ PFORM, OBL] (avec)	
R10	:	P <sup>1</sup> $\rightarrow$ H[4], NP[OBL] (sur, dedans, dans,...)	
R11	:	NP[-CLIT] $\rightarrow$ H[5] (Pierre, Marie)	
R12	:	NP $\rightarrow$ H[6] (je, tu, il, me, le,...)	
R13	:	N <sup>1</sup> [+HUM] $\rightarrow$ H[7] (garçon, fille)	
R14	:	N <sup>1</sup> [-HUM] $\rightarrow$ H[8] (cadeau, livre)	
R15	:	A <sup>1</sup> $\rightarrow$ H[9] (intelligent)	
R16	:	VP $\rightarrow$ H[10], NP[AC], NP[à, +HUM]	(présenter) — une amie au garçon
R17	:	VP $\rightarrow$ H[11], NP[AC], NP[à, -HUM]	(présenter) — une lettre au guichet

R18	: VP	→ H[12], NP[à, +HUM]	(plaire) — à la fille
R19	: VP	→ H[13], NP[à, -HUM]	(remédier) — à la douleur
R20	: VP	→ H[14], NP[à, +HUM, CLBL-]	(penser) — à lui; *lui penser
R21	: VP	→ H[15], NP[à, -HUM]	(penser) — à la promenade; y penser
R22	: VP	→ H[16], NP[AC]	(acheter) — le cadeau
R23	: VP	→ H[17], NP[de]	(raffoler) — de cela
R24	: VP	→ H[18], NP[AC], NP[de]	(convaincre) — Pierre de son erreur
R25	: VP	→ H[19], NP[AC], PP	(poser) — un livre sur la table
R26	: VP	→ H[20], AP	(être) — intelligent
R27	: VP	→ H[21], S[COMP que]	(dire) — que Pierre achète un livre
R28	: VP	→ H[22], VP[INF]	(vouloir) — partir demain

### 2.2.5. Métrarègles

M1	Inversion pronominale		
	VP	→	W
		⇓	
	VP[+INV]	→	W
M2	Placement clitique		
	VP	→	W, NP[AC]
		⇓	
	VP	→	W, NP[AC, +CLIT, PFORM de]
M3	Termination SLASH		
	X	→	W, X <sup>2</sup>
		⇓	
	X	→	W, X <sup>2</sup> [+NUL]

### 2.2.6. Stipulations LP

LP1	: [SUBCAT $\alpha$ ]	<	[~SUBCAT, [-CLIT V ~CLIT]]
LP2	: [+CLIT]	<	[SUBCAT $\alpha$ ]
LP3	: [+CLIT, +REF]	<	[+CLIT, -REF]
LP4	: [+CLIT, $\alpha$ ]	<	[+CLIT, $\alpha$ ]
	où $\alpha$ est dans {+HUM ou PFORM de}		

- LP5 : [+CLIT, NP, ~PFORM, -REF] < [+CLIT, NP,  
PFORM<sub>α</sub>, -REF]
- LP6 : [+CLIT, -REF, PFORM à] < [+CLIT, -REF,  
PFORM de]
- LP7 : [NOM] < [VP]

### 2.2.7. Entrées lexicales

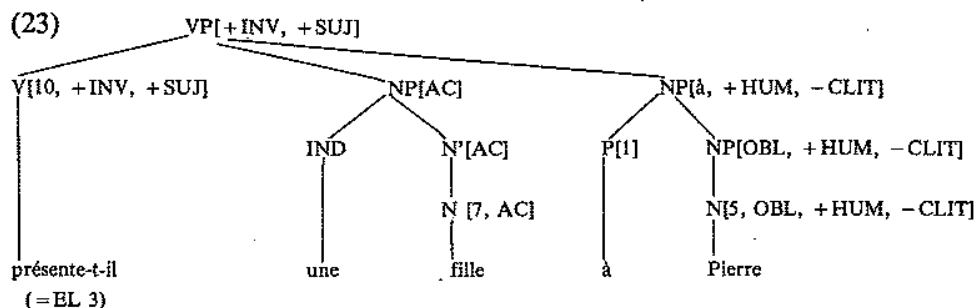
- EL 1 : <présente, [[-N], [+V], [BAR 0], [SUBCAT 10], -INV,  
FIN, PRES, AGR NP[NOM, PER 3, -PL, [GEN]], ~/]>
- EL 2 : <présentent, [[-N], [+V], [BAR 0], [SUBCAT 10], -INV,  
FIN, PRES, AGR NP[NOM, PER 3, +PL, [GEN]], ~/]>
- EL 3 : <présente-t-il, [[-N], [+V], [BAR 0], [SUBCAT 10],  
+INV, FIN, PRES, AGR NP[NOM, PER 3, -PL, MASC],  
~/]>
- EL 4 : <présente-t-il, [[-N], [+V], [BAR 0], [SUBCAT 10],  
+INV, FIN, +SUJ, PRES, AGR NP[NOM, PER 3, -PL,  
MASC],/NP[OBL, PER 3, -PL, MASC, ~Q]]>
- EL 5 : <tu, [+N, -V, [BAR 0], [SUBCAT 6], PRON, PER 2,  
-PL, GEN, +HUM, -CLIT, +REF, ~PFORM, ~WH,  
NOM, ~/]>
- EL 6 : <tu, [+N, -V, [BAR 0], [SUBCAT 6], PRON, PER 2,  
-PL, GEN, +HUM, -CLIT, +REF, ~PFORM, ~WH,  
NOM, /NP[OBL, PER 2, -PL, GEN, ~Q]]>
- EL 7 : <toi, [+N, -V, [BAR 0], [SUBCAT 6], PRON, PER 2,  
-PL, GEN, +HUM, -CLIT, +REF, ~PFORM, ~WH,  
OBL, ~/]>
- EL 8 : <te, [+N, -V, [BAR 0], [SUBCAT 6], PRON, PER 2,  
-PL, GEN, +HUM, CLBL+, +CLIT, +REF,  
~PFORM, ~WH, AC, ~/]>
- EL 9 : <te, [+N, -V, [BAR 0], [SUBCAT 6], PRON, PER 2,  
-PL, GEN, +HUM, CLBL+, +CLIT, +REF,  
[PFORM<sub>à</sub>], ~WH, ~CAS, ~/]>
- EL 10 : <te, [+N, -V, [BAR 0], [SUBCAT 6], PRON, PER 2,  
-PL, GEN, +HUM, CLBL+, +CLIT, +REF,  
~PFORM, ~WH, AC, /NP[OBL, PER 2, -PL, GEN,  
~Q]]>
- EL 11 : <te, [+N, -V, [BAR 0], [SUBCAT 6], PRON, PER 2, -PL,  
GEN, +HUM, CLBL+, +CLIT, +REF, [PFORM<sub>à</sub>],  
~WH, ~CAS, /NP[OBL, PER 2, -PL, GEN, ~Q]]>
- EL 12 : <l(e), [+N, -V, [BAR 0], [SUBCAT 6], PRON, PER 3,  
-PL, MASC, CLBL+, +CLIT, -REF, ~PFORM,  
~WH, AC, ~/]>
- EL 13 : <le, [+N, -V, [BAR 0], [SUBCAT 6], PRON, PER 3,  
-PL, MASC, CLBL+, +CLIT, -REF, ~PFORM,  
~WH, AC, /NP[PER 3, -PL, MASC, ~Q]]>
- EL 14 : <lui, [+N, -V, [BAR 0], [SUBCAT 6], PRON, PER 3,  
-PL, [GEN], +HUM, CLBL+, +CLIT, -REF,  
PFORM<sub>à</sub>, ~WH, ~CAS, ~/]>

- EL 15 : <y, [+N, -V, [BAR 0], [SUBCAT 6], PRON, PER 3, [PL], [GEN], -HUM, CLBL+, +CLIT, -REF, PFORMa, ~WH, ~CAS, ~/]>
- EL 16 : <l(e), [-N, +V, [BAR 0], [SUBCATn], PROS, +CLIT, -REF, ~PFORM, ~WH, ~CAS/S]>
- EL 17 : <l(e), [+N, +V, [BAR 0], [SUBCATn], PROAP, +CLIT, -REF, ~PFORM, ~WH, ~CAS, /AP]>
- EL 18 : <l(e), [-N, +V, [BAR 0], [SUBCATn], PROV, +CLIT, -REF, ~PFORM, ~WH, ~CAS, /VP[INF]]>
- EL 19 : <quel, [[DET INT], [BAR 0], +Q]>
- EL 20 : <e, [PX[+NUL/PX, +Q]]>

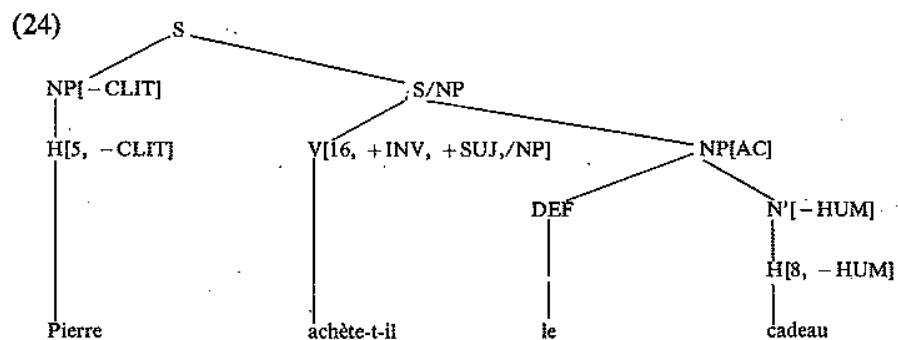
### 2.3. Commentaires

On envisagera, dans l'ordre, les trois points visés dans le champ observationnel du paragraphe 2.1. : l'inversion pronominale, les clitiques argumentaux et leur emplacement, les constructions dans la position topique et leurs relations avec les éléments de reprise dans la phrase.

L'inversion pronominale des exemples (1) et (2) (§ 2.1.) est obtenue moyennant la métarègle M1 qui opère sur les règles ID R16 à R28. On obtient ainsi, par exemple, l'arbre (23) par application de M1 sur R16 et intervention des FCR 23 et 24 :



Par utilisation de R3 on peut obtenir (24) (où *achète-t-il* est une entrée de type EL 4 ; sur l'utilisation du trait SLASH, cf. ci-dessous).



Le champ observationnel visé au paragraphe 2.1. ne concerne que les clitiques associés à des arguments verbaux. Les trois questions suivantes doivent être abordées :



- (a) l'association systématique des arguments clitiques avec les arguments non clitiques (cf. (19)) ;
- (b) l'ordre de l'emplacement des clitiques, qui diffère de l'ordre des constituants non clitiques (cf. (9) à (13)) ;
- (c) la représentation phonique qui doit être associée à des pronoms qui dénotent une même personne et qui va différer selon le constituant de la phrase qu'ils expriment (cf. (15)).

La solution ici proposée se fonde sur le formalisme des règles ID, dissociées des relations d'ordre exprimées dans les stipulations LP, et sur l'utilisation des FCR et FSD dans la constitution des catégories. En lignes générales, elle consiste à proposer, pour les entrées verbales, des règles lexicales ID très peu spécifiées. Dans les filles de ces règles, il sera possible d'instancier le trait  $\langle \text{CLIT} ; \pm \rangle$ , dont la présence, combinée avec la présence d'autres traits, déclenchera, grâce aux FCR et aux FSD, une série d'effets dans la conformation des catégories. Les stipulations LP vont opérer sur les catégories ainsi constituées afin d'obtenir les emplacements adéquats.

A l'exception des catégories [CLBL -] (cf. R20) et des entrées verbales (cf. FCR 28), il est possible d'instancier [CLIT] sur les filles des règles ID R16 à R28. Par ailleurs, M2 permettra de préparer la catégorie associée à *en* complément d'objet direct (cf. (14 C')). Ce sont les deux mécanismes qui permettent d'exprimer une association systématique des arguments clitiques avec les arguments non clitiques.

Sauf exigence contraire (cf. FSD 8), une catégorie NP complément d'objet direct ou indirect est [CLBL +]. Si par ailleurs elle est [- CLIT], aucun pronom, par FCR 10, ne sera un symbole terminal adéquat ; on évite ainsi (25) (avec *à toi* comme complément d'objet indirect), et par R7 et R11 on obtient (26) (et par R7 et R6, (27)) :

- (25) \*Pierre présente une amie à toi
- (26) Pierre présente un ami à Marie
- (27) Pierre présente un ami à la fille

L'emplacement des clitiques — et plus généralement, de toutes les catégories — est réglé par les stipulations LP, qui vont spécifier les relations d'ordre autorisées pour les catégories instanciées dans les arbres locaux. Avant de présenter ces stipulations, il est donc nécessaire d'observer comment les FCRs et les FSDs permettent de constituer ces catégories.

Par le FCR 4, les catégories clitiques ([+ CLIT]) auront une valeur de PROF. Par les FCRs 5 à 8, cette valeur sera déterminée par le type de catégorie où apparaît [+ CLIT], de telle manière que, par exemple, la catégorie NP fille dans R22 sera PRON et la catégorie AP fille de R26 sera PROA. Les clitiques pronominaux auront une valeur pour REF (cf. FCR 9), et cette valeur sera + pour les pronoms de première et deuxième personnes (cf. FCRs 1 et 2) et pour les réflexifs (cf. FCR 12). En principe, les catégories ne sont pas marquées pour CAS (cf. FSD 3), sauf exigence contraire. Cette exigence apparaît soit dans les règles ID, où on incorpore AC et OBL, soit dans les FCRs 17 et 26, qui, respectivement, exigeront l'incorporation de OBL dans le complément topique, et NOM dans la catégorie NP sujet. FCR 3 exige [- CLIT] dans les catégories [NOM] et dans les catégories [OBL].

Les stipulations LP vont spécifier les relations d'ordre entre les catégories. LP1 exige que les catégories qui ont une valeur pour SUBCAT précèdent celles qui n'en ont pas, pourvu soit que celles-ci ne possèdent pas de valeur pour CLIT, soit qu'elles aient comme valeur [-CLIT]. Grâce à LP1, le déterminant sera placé devant le nom, la préposition devant son complément nominal, le complémenteur devant la phrase, et l'entrée verbale devant les compléments argumentaux non clitiques. LP2 exige que les catégories clitiques soient placées devant l'entrée verbale. LP7 exige que le NP sujet soit placé devant le syntagme verbal ; elle s'appliquera à l'arbre local projection de R1, mais non à celui qui est une projection de R2, car le NP en position topique n'est pas NOM.

Les autres LP (LP3 à LP6) concernent les catégories clitiques et sont destinées à rendre compte de la matrice de (9). Parmi les interdictions de (9), il y en a de deux types : celles qui découlent de l'organisation générale de la description et, tout particulièrement, de la manière d'associer les catégories clitiques aux arguments verbaux, et celles qui exigent des formulations supplémentaires. La transition *te (ac) se (ac)* illustre une interdiction du premier type : étant donné qu'il n'y a qu'une fille NP[AC] par entrée verbale, cette catégorie deviendra dans les arbres locaux soit *te (ac)* soit *se (ac)*, mais aucune projection légale ne présentera la transition entre les deux. Dans la matrice (29), ces interdictions seront marquées  $\emptyset$ . Les LP 3 à 6 sont destinées à rendre compte des autres interdictions de (9).

Il est possible de faire une partition des formes clitiques ( { *me, te, se, le, lui, y, en* } ) en deux ensembles { *me, te, se* }<sub>A</sub> et { *le, lui, y, en* }<sub>B</sub>. Les régularités suivantes peuvent être formulées :

- (28) a Un clitique de l'ensemble A précède un clitique de l'ensemble B.  
 b *Lui* ne doit pas être précédé d'un clitique de l'ensemble A.  
 c *Le* doit précéder *lui* et *y*.  
 d *Lui* et *y* précèdent *en*.  
 e *En* ne précède pas *en*.

LP3 permet d'exprimer une partie des observations couvertes par (28a), celles qui concernent l'impossibilité de transition entre un clitique de l'ensemble B et un clitique de l'ensemble A.

LP4 rend compte des cas de (28a) non couverts par LP3 (en l'occurrence, l'impossibilité de transition entre deux clitiques de l'ensemble A), de (28b) et de (28c). LP4 profite en effet du formalisme des LP (cf. au § 1 *Acceptabilité LP*) permettant d'interdire la transition entre deux catégories identiques. Par HFC, *lui* hérite de [HUM +] et les clitiques de l'ensemble A sont tous [HUM +]. Le même principe va s'appliquer à l'interdiction de transition entre clitiques de l'ensemble A, qui sont [HUM +], et entre deux clitiques *en*, qui sont [PFORM *de*] (l'un a acquis ce trait par M2, l'autre l'hérite de l'entrée verbale).

LP5 exprime (28c). En effet, d'après FD6, les catégories nominales ne présentent pas de valeur pour PFORM, sauf exigence contraire. Cette exigence apparaît dans les NP filles des entrées verbales autres que le

complément d'objet direct, et est héritée par les catégories clitiques. Les clitiques [AC] vont ainsi précéder ceux qui sont [PFORM  $\alpha$ ].

LP6 exprime enfin (28d), car *lui* et *y* sont [PFORM $\delta$ ].

Il est ainsi possible de reprendre (9) dans (29) et d'y noter, pour chaque intersection marquée \* dans (9), comment la grammaire en rend compte :  $\emptyset$  dans (29) indique que le schéma général d'association entre formes clitiques et non clitiques interdit la transition en question (cf. ci-dessus) ; dans les autres cas d'interdiction, un numéro renvoie à la stipulation LP correspondante. Les situations notées G ne sont interdites par aucune stipulation LP et sont spécifiées par la grammaire<sup>5</sup> :

(29)

	se(ac)	se(dat)	te(ac)	te(dat)	le	lui	y	en(ac)	en(de)
se(ac)	$\emptyset$	4	$\emptyset$	4	$\emptyset$	4	G	$\emptyset$	G
se(dat)	4	$\emptyset$	4	$\emptyset$	G	$\emptyset$	$\emptyset$	G	G
te(ac)	$\emptyset$	4	$\emptyset$	4	$\emptyset$	4	G	$\emptyset$	G
te(dat)	4	$\emptyset$	4	$\emptyset$	G	$\emptyset$	$\emptyset$	G	G
le	$\emptyset$	3	$\emptyset$	3	$\emptyset$	G	G	$\emptyset$	G
lui	3	$\emptyset$	3	$\emptyset$	5	$\emptyset$	$\emptyset$	G	G
y	3	$\emptyset$	3	$\emptyset$	5	$\emptyset$	$\emptyset$	G	$\emptyset?$
en(ac)	$\emptyset$	6	$\emptyset$	6	$\emptyset$	6	6	$\emptyset$	4
en(de)	3	3	3	3	5	6	6	4	$\emptyset$

La troisième question relative aux clitiques et à leur emplacement, à savoir celle concernant leur représentation phonique (cf. (15)), est réglée, dans la solution ici proposée, moyennant les valeurs de CAS exigées par la grammaire (cf. ci-dessus) et par les catégories des EL (cf. § 2.2.7.). Etant donné l'exigence sur les arbres locaux terminés (cf. § 1. *in fine*) un pronom OBL (cf. EL 7) ne peut être symbole terminal qu'en position topique ou comme complément de préposition ; un pronom NOM (cf. EL 5 et EL 6) ne peut être que sujet. Or, ce sont les positions d'où les clitiques sont exclus : cf. les autres entrées du pronom de deuxième personne (EL 8 à EL 11).

Il reste maintenant à considérer comment la grammaire exprime les rapports formulés dans (19) et (20) à propos des constructions dans la position topique et leurs relations avec les proformes de reprise.

Le formalisme de GPSG permet de rendre compte des relations entre éléments arbitrairement éloignés sans avoir recours à des transformations. Les éléments essentiels qui le permettent sont :

- (a) le trait SLASH, à valeur catégorielle, et le principe FFP ;
- (b) la règle ID R3 et la métarègle M3 ;
- (c) le FCR 25, les FSD 2 et 7 et le trait schématique :

(30)    /[PX]

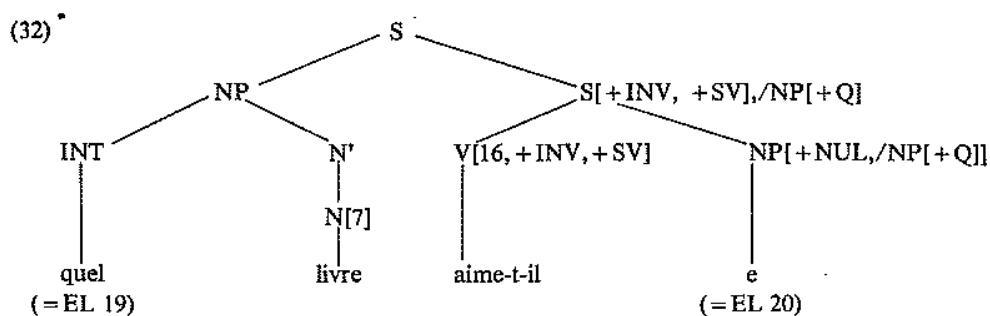
Le principe FFP va permettre de transmettre dans l'arbre de la mère à une de ses filles l'information sur l'existence d'une catégorie avec le trait SLASH. La règle ID R3 incorpore une variable : la valeur de SLASH est celle de la catégorie en position topique. L'arbre local qui en est une projection aura SLASH comme trait instancié, mais l'arbre local qui est imbriqué dans lui et dont la mère est S/NP, aura SLASH comme

trait instancié. C'est à partir de cet arbre-ci que le principe FFP agira, transmettant le trait SLASH avec sa valeur, qui correspond à la catégorie dans la position topique.

La métarègle M3 incorpore le trait [+NUL] à une catégorie fille d'une règle ID lexicale. On obtient ainsi, par exemple :

(31) R'22 VP → H[16], NP[AC, +NUL]

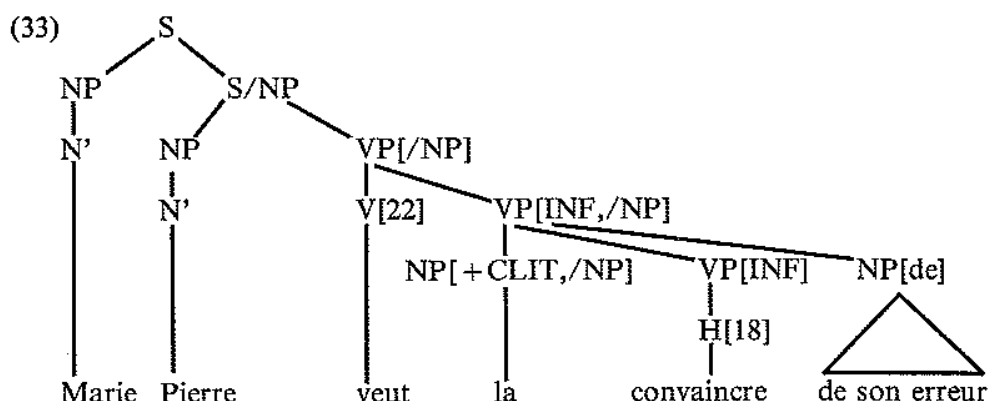
Grâce au FSD 2, il n'y aura pas d'autres catégories [+NUL] que celles licenciées par les règles ID obtenues à partir de M3 ; FCR 27 y exigera le trait SLASH. Dans le lexique, on a EL 20 qui correspond à un élément sans spécification phonique et qui incorpore (30). On obtient ainsi l'arbre qui correspond à (22a).



Dans les entrées clitiques et dans les entrées verbales [+INV], PX de (30) prendra différentes valeurs. Dans les entrées verbales, XP = NP (cf. EL 4) ; on spécifie ainsi l'arbre qui correspond à (21 b). Dans les entrées clitiques, si le clitique est [PRON], PX = NP ; mais on a aussi PX = AP, PX = NP[à], PX = VP et PX = S, ce qui rend compte de (4) à (8) (dans l'échantillon des entrées lexicales du paragraphe 2.2.7., on a incorporé celles qui correspondent à NP(EL 4), à AP(EL 17), à VP(EL 18) et à S(EL 16). Lorsque (30) est incorporé à une entrée lexicale avec spécification phonique, la catégorie valeur de SLASH est [~Q] ; en revanche, elle est [+Q] dans l'élément vide (cf. EL 20). On peut ainsi exprimer (19) et (20). Comme dans le lexique il n'y a pas de clitique avec une valeur de PROF autre que NP mais sans trait SLASH, et ceci contrairement aux clitiques PRON (cf. par exemple les EL 8 et 9), la grammaire spécifie (2) et (17), (32a) et (32b), mais elle ne spécifie pas (32c) :

- (32) (a) Marie, Pierre la regarde  
 (b) Pierre la regarde  
 (c) \*Pierre l'est

L'information transmise par SLASH et le principe FFP se propage dans les phrases enchâssées, de telle manière qu'il est également possible d'exprimer (19) et (29) si la catégorie de la position topique est arbitrairement éloignée de la proforme qui doit lui être associée :



### 3. TRADUCTION SÉMANTIQUE ET PARAPHRASE

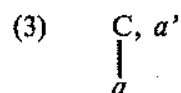
L'appareil syntaxique de *GPSG* résumé au paragraphe 1 permet de spécifier un sous-ensemble intéressant d'énoncés, dont

- (1) Pierre regarde Marie
- (2) Marie, Pierre la regarde

La sémantique de *GPSG* va permettre de formaliser la relation de paraphrase entre (1) et (2).

*GPSG* profite, dans sa sémantique, de l'apport conceptuel de la grammaire de Montague. Comme chez Montague, la sémantique de *GPSG* est compositionnelle et vériconditionnelle. Un arbre syntaxique terminé sera traduit dans une expression en logique intensionnelle et celle-ci sera interprétée dans un univers d'interprétation.

On sait (cf. § 1) que les symboles terminaux des arbres terminés sont des entrées lexicales, avec leurs traductions sémantiques. On peut ainsi se donner les arbres locaux terminés interprétés, qui schématiquement sont représentés comme dans (3) :



On a  $a$  = entrée lexicale,  $a'$  = constante non logique, traduction sémantique de l'entrée  $a$ ,  $C$  = catégorie mère (cf. § 1 (29)).

Les arbres locaux terminés constituent l'entrée du schéma d'interprétation sémantique : le rôle de celui-ci consiste à assigner une traduction sémantique à chaque catégorie mère dans un arbre local à partir des traductions sémantiques assignées aux catégories filles. On obtient ainsi un arbre interprété ; la traduction assignée à la catégorie racine  $S$  de l'arbre associé à une phrase  $P$  sera la traduction en logique intensionnelle de  $P$  et l'interprétation dans l'univers sera faite à partir de cette traduction.

Le schéma d'interprétation sémantique opère essentiellement en fonction des éléments suivants :

- (a) les types sémantiques assignés à chaque catégorie syntaxique ;
- (b) l'opération d'application fonctionnelle ;
- (c) l'information contenue dans les catégories des arbres locaux.

Les types sémantiques sont les catégories syntaxiques de la LI. Les types de base sont *e* et *t*, les autres étant obtenus inductivement à partir d'eux :

- (a) Si *a* est un type,  $\langle s, a \rangle$  est aussi un type
- (b) Si *a* et *b* sont des types,  $\langle a, b \rangle$  est aussi un type.

L'assignation des types sémantiques aux catégories syntaxiques ne se fait pas en *GPSG* par règle mais par spécification individuelle. Les types des catégories non lexicales dont on aura besoin dans les exemples ici traités sont les suivants :

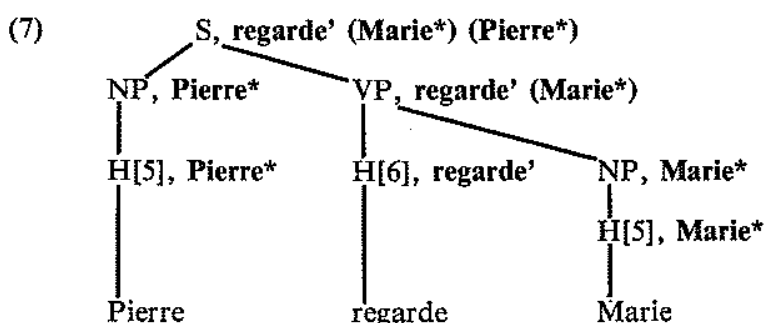
(5) Catégories	Type	Notation abrégée
S	$\langle s, t \rangle$	$\langle S \rangle$
NP	$\langle s, \langle \langle e, t \rangle, t \rangle \rangle$	$\langle NP \rangle$
VP	$\langle \langle s, \langle \langle e, t \rangle, t \rangle \rangle, \langle s, t \rangle \rangle$ = $\langle NP, S \rangle$	$\langle VP \rangle$

Le type assigné à une tête lexicale sera une fonction qui va associer la ou les catégories sœurs de la tête lexicale à la catégorie mère. Par exemple, le type assigné à H[16] (cf. règle ID 22) sera  $\langle NP, VP \rangle$ .

La construction des traductions sémantiques se fait par application fonctionnelle ; le principe général de cette opération se résume ainsi :

- (6) Si  $\alpha$  est une expression de type  $\langle b, a \rangle$  et  $\beta$  une expression de type  $\langle b \rangle$ ,  $\alpha(\beta)$  est une expression de type  $\langle a \rangle$ .

Un des intérêts de la sémantique de *GPSG* est, semble-t-il, que l'opération (6) est incorporée au schéma d'interprétation sémantique : celui-ci consiste essentiellement à assigner une expression en LI à la catégorie mère par application de (6) aux expressions en LI associées aux catégories filles ; cette assignation se fait en tenant compte des types sémantiques assignés à la mère et aux filles. La traduction n'est pas ainsi explicitée règle par règle, mais construite en fonction de principes généraux. L'arbre interprété (7) est ainsi<sup>6</sup> celui qui doit être associé à (1)



La possibilité d'avoir accès à l'information portée par les catégories dans les arbres locaux constitue une autre caractéristique importante de la sémantique de *GPSG*. Elle va permettre de moduler l'application du schéma d'interprétation sémantique en fonction des traits instantiés, et, en tant que tels, absents des règles ID correspondantes. Cette possibilité sera notamment utilisée pour les catégories avec SLASH instancié. Le schéma d'interprétation sémantique incorpore en effet, à leur égard, un principe, ici condensé conceptuellement dans (8), qui permet de moduler l'application de (6) :

- (8) La traduction sémantique d'une catégorie mère avec SLASH instancié est une formule lambda dans laquelle la variable lambda (=  $P$ ) est liée. L'application fonctionnelle sur les catégories filles se fera sur une traduction de la catégorie fille avec SLASH instancié qui sera  $a'(P)$ .

Soit l'arbre local (9) (par FFP, /NP est instancié et dans la mère et dans une catégorie fille ; entre parenthèses, les expressions sémantiques qui sont la traduction des catégories syntaxiques filles correspondantes).

- (9) S/NP  
       NP( $b'$ )  
       VP/NP( $a'(P)$ )

Par application de (8), la traduction de la catégorie mère de (9) sera :

- (10)  $\lambda P[a'(P) (b')]$

Les arbres locaux qui sont une projection de R3 présentent une catégorie fille (en l'occurrence S/NP) avec SLASH licencié (et non instancié, comme dans (9)). Ceci permet d'introduire une nouvelle convention portant sur la traduction de la catégorie mère de (11). Elle est exprimée en (12) et la valeur sémantique de EXT explicitée dans (13) :

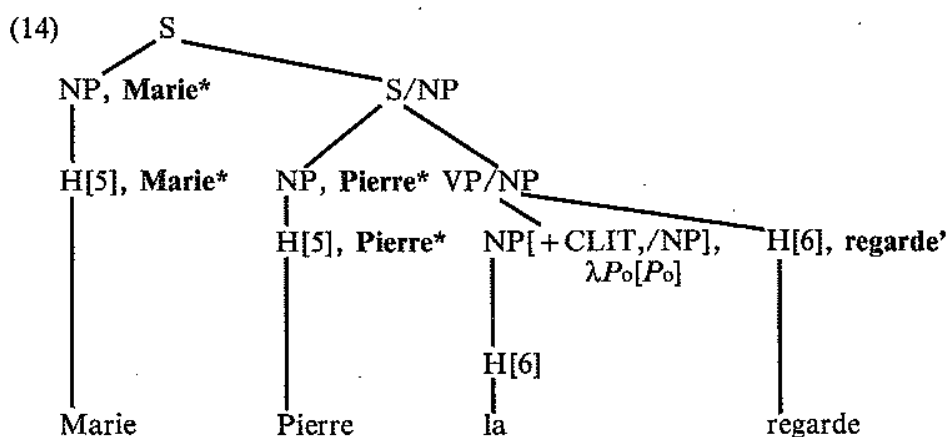
- (11) S  
       NP( $b'$ )  
       S/NP( $a'$ )

- (12) EXT( $a'$ ) ( $b'$ )

- (13) EXT =  $\lambda v^{VP} \lambda P [\neg P \{ \lambda x [v^{VP}(x^*)] \} ]$

Il reste maintenant à introduire la traduction sémantique de l'élément nul (cf. EL 20), et plus généralement, des entrées clitiques pronominales avec trait SLASH (cf. par exemple EL 13) :  $\lambda P_0[P_0]$ .

A partir de (8), (12) et (13) il est possible d'obtenir l'arbre interprété qui doit être associé à (2) :



Dans (14), les traductions sémantiques associées à VP/NP, S/NP et S sont, respectivement :

- (14) (a)  $\lambda P_1[\text{regarde}'(\lambda P_0[P_0] (P_1))]$   
 $\iff \lambda P[\text{regarde}'(P_1)]$   
 (b)  $\lambda P_1[\lambda P_0[\text{regarde}'(P_0)] (P_1) (\text{Pierre}^*)]$   
 $\iff \lambda P_1[\text{regarde}'(P_1) (\text{Pierre}^*)]$   
 (c)  $\lambda v^{VP} \lambda P[ P \{ \lambda x[v^{VP}(x^*)] \} ] (\lambda P_1[\text{regarde}'(P_1) (\text{Pierre}^*)]) (\text{Marie}^*)$

Les équivalences de (14a) — traduction de VP/NP — et de (14b) — traduction de S/NP — sont obtenues par réduction lambda. La formule de (14c) correspond à la catégorie S racine et est la traduction en LI de (2). Par réduction lambda et par simplification de l'extension-intension on peut ramener (14c) — réintroduit comme (15a) — à (15f), c'est-à-dire à la traduction associée à (1) (cf. l'arbre terminé (7)) :

- (15) (a)  $\lambda v^{VP} \lambda P[ P \{ \lambda x[v^{VP}(x^*)] \} ] (\lambda P_1[\text{regarde}'(P_1) (\text{Pierre}^*)]) (\text{Marie}^*)$   
 (b)  $\lambda v^{VP}[ (\text{Marie}^*) \{ \lambda x[v^{VP}(x^*)] \} ] (\lambda P_1[\text{regarde}'(P_1) (\text{Pierre}^*)])$   
 (c)  $\wedge (\text{Marie}^*) \{ \lambda x[\lambda P_1[\text{regarde}'(P_1) (\text{Pierre}^*)]] (x^*) \}$   
 (d)  $\wedge (\text{Marie}^*) \{ \lambda x[\text{regarde}'(x^*) (\text{Pierre}^*)] \}$   
 (e)  $(\text{Marie}^*) \lambda x[\text{regarde}'(x^*) (\text{Pierre}^*)]$   
 (f)  $\text{regarde}' (\text{Marie}^*) (\text{Pierre}^*)$

D'une manière élégante, GPSG réussit ainsi à formaliser la relation paraphrastique entre (1) et (2). Il le fait sans avoir recours à une structure syntaxique abstraite ou profonde et, dans ce cas, sans avoir recours à des postulats de signification. Parce qu'il incorpore des sémantiques calculables, c'est-à-dire dont on connaît les propriétés formelles, il est possible d'associer à des arbres syntaxiques différents des expressions sémantiques qui, certes, sont différentes mais qui, grâce aux propriétés formelles du langage dans lesquelles elles sont exprimées, sont équivalentes.



## 4. EXTENSIONS ET PROBLÈMES

La solution proposée dans les paragraphes qui précèdent exploite trois caractéristiques remarquables du formalisme de *GPSG* :

- (a) la dissociation des relations d'ordre de la relation de dominance dans les règles ID ;
- (b) la possibilité d'exprimer les rapports à distance moyennant le trait SLASH et le principe FFP ;
- (c) la sémantique calculable associée aux arbres syntaxiques.

Il semble bien que la solution proposée aux paragraphes 2 et 3 rend compte de manière systématique d'un langage relativement intéressant du français. Mais il doit être non moins clair que tous les problèmes concernant les clitiques et les constructions topiques ne sont pas pour autant épuisés. Et parmi les problèmes qui restent, il y en a qu'il est impossible de présenter comme nécessitant simplement d'introduire de simples ajustements à l'analyse précédente. On esquissera une rapide présentation ici des problèmes non traités dans les paragraphes 2 et 3, en essayant d'évaluer les possibilités du modèle pour en rendre compte.

Les relations paraphrastiques n'existent pas seulement entre constructions nominales en position topique et clitiques pronominaux. On les constate aussi avec les autres pro-formes :

- (1) a Intelligente, Marie l'est
- a' Marie est intelligente
- b A la poste, Pierre y va souvent
- b' Pierre va souvent à la poste
- c Apprendre les mathématiques, Pierre le souhaite
- c' Pierre souhaite apprendre les mathématiques
- d Qu'il est déjà tard, Pierre le sait
- d' Pierre sait qu'il est déjà tard.

Il ne paraît pas impossible de traiter ces cas de manière tout à fait analogue aux cas traités dans le paragraphe 3, où seuls les clitiques pronominaux étaient en jeu. La modulation (8) du paragraphe 3 peut se généraliser en typant les variables sur lesquelles porte l'opérateur lambda<sup>7</sup>. De la même manière, il sera nécessaire de généraliser la définition (13) du paragraphe 3, mais ici le problème se présente comme plus délicat car, au moins pour (14c), la construction topique ne semble pas extensionnelle.

Les clitiques peuvent précéder les formes d'infinitif et les formes participiales utilisées adverbiallement. La généralisation à ces cas de l'analyse précédente est très directe en *GPSG* ; on aura simplement des VP avec une valeur de VFORM autre que FIN ; pour le reste, la solution précédente devrait s'appliquer sans trop de difficulté. En revanche, un réel problème pour *GPSG* est l'emplacement des clitiques argumentaux à la gauche de l'auxiliaire dans les formes verbales composées :

- (1) a Marie les lui a donnés
- b Pierre lui en a parlé

Ce même type de difficulté existera aussi pour les clitiques qui correspondent à des extractions d'un NP (cf. (3)) ou à des arguments adjectivaux (cf. (4)) :

- (3) Pierre en a pris un
- (4) Pierre en est conscient

Dans les exemples (2) à (4) il est nécessaire de transférer l'information entre catégories filles d'arbres locaux différents : il n'est pas impossible que SLASH et FFP puissent rendre ce service.

Il n'est pas vrai que la position topique doit être occupée par une construction unique :

- (5) a Pierre, Marie, il la regarde
- b A la poste, Pierre, il y a va tous les jours.

Ces cas restent en dehors de la portée de l'analyse ici présentée. Le formalisme utilisé n'admet qu'un trait SLASH qui a, comme valeur, une seule catégorie. Mais il n'est pas unimaginable d'élargir ce formalisme en le rendant moins restrictif ; par ailleurs, l'opérateur utilisé par GPSG pour la coordination pourrait aussi rendre service pour concaténer les constructions dans la position topique<sup>8</sup>.

Mais il y a des cas qui semblent constituer un obstacle infranchissable pour le modèle : ce sont les clitiques non argumentaux. GPSG fait une prévision importante et explicite concernant la linéarité dans les langues naturelles :

- (6) By requiring that linear precedence statements be satisfied in the same way as other principles of local admissibility, i.e. locally for each mother-and-daughters substructure rather than globally over whole trees at once, we predict that ordering constraints in natural languages are limited to sisters — for example, that order verb and object cannot depend on whether there is an adverb in a nearby AP — without having to stipulate this as an extra condition on linearization rules.

(p. 105)

L'emplacement des clitiques en français ne confirme pas la prévision de (6). Même si la ligne de clivage entre arguments verbaux et compléments verbaux est loin d'être transparente, il est relativement facile de se donner des cas où, de manière claire, des formes clitiques n'expriment pas des arguments verbaux. Dans (7), *se* est un datif d'intérêt et non un complément d'objet indirect de *farcir* ; dans (8), *y* est un complément de lieu, du même type que *à Paris*, dans (9) :

- (7) Les contraintes des règles, Jacques se les est farcies à fond
- (8) Jacques y a consulté un bon nombre de spécialistes
- (9) A Paris, Jacques a consulté un bon nombre de spécialistes

Or, ce qui paraît remarquable concernant l'ordre d'emplacement des clitiques en français, c'est que les contraintes de transition sont les mêmes entre clitiques non argumentaux :

- (10) \*Les contraintes des règles, Jacques les se est farcies à fond.
- (11) a Jacques y en a déjà consulté (plusieurs)  
 a' \*Jacques en y a déjà consulté (plusieurs)  
 b Jacques les y a déjà consultés  
 b' \*Jacques y les a déjà consultés

Ces phénomènes semblent bien rester en dehors des limites de principe du formalisme de *GPSG* : en effet, pour rendre compte des cas observés de (7) à (11), il est nécessaire d'introduire des stipulations entre catégories filles d'arbres locaux différents, ce que les stipulations LP ne peuvent pas faire ; la prévision (6) dans un sous-langage bien caractérisé, se voit ainsi invalidée.

*GPSG*, avec son effort de clarté, de rigueur et de formalisation au sens fort, est sans doute un point de référence dans l'histoire de la linguistique. Point d'achèvement ou point de commencement d'une nouvelle génération de grammaires ? On aimerait penser que la réponse se trouve quelque part entre les deux termes de la disjonction. En tout cas, dans ce numéro de *Lexique* consacré à la paraphrase, il semblait judicieux de présenter un modèle dont les échecs — parce que contrôlés — préparent les solutions linguistiques du troisième millénaire. Avec une syntaxe simplifiée mais un lexique alourdi, et une sémantique formalisée et calculable, *GPSG* permet d'exprimer les relations entre conditions de vérité associées à des structures syntaxiques différentes. Ceci ne résout certes pas tout ce que la notation de paraphrase est censée couvrir, mais il serait raisonnable de souhaiter que des notions de paraphrase plus ambitieuses s'expriment, elles aussi, dans un formalisme calculable et incorporent l'expression componentielle des conditions de vérité.

GABRIEL G. BÈS  
*Université de Clermont II*  
*(CILN — Formation Doctorale*  
*Linguistique et Informatique)*  
*SILEX, UA 382 du C.N.R.S.*

## NOTES

1. Cf. Gazdar (1982), Pullum & Gazdar (1982), Bresnan (1982).
  2. Par exemple, d'après les définitions de CAP et en tenant compte des FCRs 12, 13 et FSDs 4 et 10, *GPSG* admet l'accord en anglais entre verbe et complément d'objet direct (cf. Bès & Baschung (1985)).
  3. On utilise les mêmes acronymes qu'en anglais. Dans les exemples, on n'indique que les traits de chaque catégorie nécessaires à la discussion. Abréviations utilisées : S = [+V, -N, BAR 2, +SUJ] ; VP = [+V, -N, BAR 2, -SUJ] ; NP = [-V, +N, BAR 2].
- Pour une présentation formelle des matières syntaxiques du paragraphe 1 et des matières sémantiques du paragraphe 3, le lecteur est invité à se reporter à *GPSG*.
- Les conventions de notation de cet article sont valables également pour l'article

de Baschung, mais elles diffèrent de celles du travail de Fradin. Dans celui-ci, « GSG » = GPSG, « traits de restrictions de cooccurrence » = FCR, les valeurs de BARRE sont notées en exposant, les règles ID sont notées autrement (cf. Fradin, § 3.), et les interprétations sémantiques sont proposées règle par règle et non moyennant le schéma d'interprétation utilisé dans cet article au paragraphe 3.

4. Dans (9), *te(ac)* couvre *me(ac)*, *nous(ac)* et *vous(ac)*; de manière analogue, *te(dat)*: *me(dat)*, *nous(dat)*, *vous(dat)*; *le*: *la*, *les*; *lui*: *leur*.

5. Cf. le même tableau, mais avec un traitement différent dans Baschung & alii (1987). Pour une vision d'ensemble des problèmes posés par les clitiques en français, cf. Blanche-Benveniste & al. (1984, chap. 2).

6. On note en caractères gras la traduction de constantes non logiques.

7. Cf. cette solution dans GPSG, pp. 229-236.

8. Cf. le « iterating coordination schema » de GPSG, p. 171.

## BIBLIOGRAPHIE

- BACH E. (1965), « Linguistique structurale et philosophie des sciences », *Diogène* 51, pp. 117-136.
- BASCHUNG K., BÈS G.G., CORLUIY A. & GUILLOTIN T. (1987), « Auxiliaries and Clitics in French UCG Grammar », Communication présentée à The Third Conference of the European Chapter of the Association for Computational Linguistics, Copenhagen.
- BÈS G.G. (1986), « Clíticos en francés y modelos lingüísticos », *Revista Argentina de Lingüística* V.2, n° 2, pp. 246-265.
- BÈS G.G. & BASCHUNG K. (1985), « Feasibility of a GPSG French Grammar », *Rapport PROJET ESPRIT 393 ACORD*.
- BLANCHE-BENVENISTE C., DEULOFEU J., STÉFANINI J. & Van den EYNDE K. (1984), *PRONOM ET SYNTAXE: L'approche pronominale et son application au français*, Paris, Selaf.
- BRESNAN J. ed. (1982), *The Mental Representation of Grammatical Relations*, Cambridge, Mass., The M.I.T. Press.
- GAZDAR G. (1982), « Phrase Structure Grammar », in Jacobson P. & Pullum G.K. eds, *The Nature of Syntactic Representation*, Dordrecht, Reidel, pp. 131-186.
- GAZDAR G., KLEIN E., PULLUM G. & SAG I. (1985), *Generalized Phrase Structure Grammar*, Cambridge, Mass., Harvard University Press.
- PULLUM G. & GAZDAR G. (1982), « Natural Languages and context free Languages », *Linguistics and Philosophy* 4, pp. 471-504.